

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

Nr 327

Taksonomia 22

**Klasyfikacja i analiza danych –
teoria i zastosowania**

Redaktorzy naukowci

Krzysztof Jajuga, Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2014

Redaktor Wydawnictwa: Barbara Majewska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:

www.ibuk.pl, www.ebscohost.com,

w Dolnośląskiej Bibliotece Cyfrowej www.dbc.wroc.pl,

The Central and Eastern European Online Library www.ceeol.com,

a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon

http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

Informacje o naborze artykułów i zasadach recenzowania znajdują się

na stronie internetowej Wydawnictwa

www.wydawnictwo.ue.wroc.pl

Tytuł dofinansowany ze środków Narodowego Banku Polskiego

oraz ze środków Sekcji Klasyfikacji i Analizy Danych PTS

Kopiowanie i powielanie w jakiegokolwiek formie

wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu

Wrocław 2014

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)

ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp	9
Eugeniusz Gatnar , Balance of payments statistics and external competitiveness of Poland.....	15
Andrzej Sokolowski, Magdalena Czaja , Efektywność metody k -średnich w zależności od separowalności grup.....	23
Barbara Pawelek, Józef Pocięcha, Adam Sagan , Wielosektorowa analiza ukrytych przejść w modelowaniu zagrożenia upadłością polskich przedsiębiorstw	30
Elżbieta Gołata , Zróżnicowanie procesu starzenia i struktur demograficznych w Poznaniu i aglomeracji poznańskiej na tle wybranych dużych miast Polski w latach 2002-2011.....	39
Aleksandra Łuczak, Feliks Wysocki , Ustalanie systemu wag dla cech w zagadnieniach porządkowania liniowego obiektów	49
Marek Walesiak , Wzmacnianie skali pomiaru dla danych porządkowych w statystycznej analizie wielowymiarowej	60
Paweł Lula , Identyfikacja słów i fraz kluczowych w tekstach polskojęzycznych za pomocą algorytmu <i>RAKE</i>	69
Mariusz Kubus , Propozycja modyfikacji metody złagodzonego LASSO.....	77
Andrzej Bąk, Tomasz Bartłomowicz , Wielomianowe modele logitowe wyborów dyskretnych i ich implementacja w pakiecie <i>DiscreteChoice</i> programu R.....	85
Justyna Brzezińska , Wykorzystanie modeli logarytmiczno-liniowych do analizy bezrobocia w Polsce w latach 2004-2012.....	95
Andrzej Bąk, Marcin Pelka, Aneta Rybicka , Zastosowanie pakietu <i>dcMNM</i> programu R w badaniach preferencji konsumentów wódki	104
Barbara Batóg, Jacek Batóg , Analiza stabilności klasyfikacji polskich województw według sektorowej wydajności pracy w latach 2002-2010	113
Małgorzata Markowska, Danuta Strahl , Klasyfikacja europejskiej przestrzeni regionalnej ze względu na filary inteligentnego rozwoju z wykorzystaniem referencyjnego systemu granicznego.....	121
Kamila Migdał-Najman, Krzysztof Najman , Formalna ocena jakości odwzorowania struktury grupowej na mapie Kohonena	131
Kamila Migdał-Najman, Krzysztof Najman , Graficzna ocena jakości odwzorowania struktury grupowej na mapie Kohonena	139
Beata Basiura, Anna Czapkiewicz , Badanie jakości klasyfikacji szeregów czasowych	148
Michał Trzęsiok , Wybrane metody identyfikacji obserwacji oddalonych.....	157

Grażyna Dehnel, Tomasz Klimanek , Taksonomiczne aspekty estymacji pośredniej uwzględniającej autokorelację przestrzenną w statystyce gospodarczej.....	167
Michał Bernard Pietrzak, Justyna Wilk , Odległość ekonomiczna w modelowaniu zjawisk przestrzennych z wykorzystaniem modelu grawitacji.....	177
Maciej Beręsewicz , Próba zastosowania różnych miar odległości w uogólnionym estymatorze Petersena	186
Marcin Szymkowiak, Tomasz Józefowski , Konstrukcja i praktyczne wykorzystanie estymatorów typu SPREE na przykładzie dwuwymiarowych tabel kontyngencji	195
Marcin Pelka , Klasyfikacja pojęciowa danych symbolicznych w podejściu wielomodelowym	202
Małgorzata Machowska-Szewczyk , Ocena klas w rozmytej klasyfikacji obiektów symbolicznych.....	210
Justyna Wilk , Problem wyboru liczby klas w taksonomicznej analizie danych symbolicznych.....	220
Andrzej Dudek , Metody analizy skupień w klasyfikacji markerów map Google	229
Ewa Roszkowska , Ocena ofert negocjacyjnych w słabo ustrukturyzowanych problemach negocjacyjnych z wykorzystaniem rozmytej procedury SAW	237
Marcin Szymkowiak, Marek Witkowski , Zastosowanie analizy korespondencji do badania kondycji finansowej banków spółdzielczych.....	248
Bartłomiej Jefmański , Budowa rozmytych indeksów satysfakcji klientów z zastosowaniem programu R.....	257
Karolina Bartos , Odkrywanie wzorców zachowań konsumentów za pomocą analizy koszykowej danych transakcyjnych	266
Joanna Trzęsiok , Taksonomiczna analiza krajów pod względem dzietności kobiet oraz innych czynników demograficznych	275
Beata Bal-Domańska , Próba identyfikacji większych skupisk regionalnych oraz ich konwergencja.....	285
Beata Bieszk-Stolorz, Iwona Markowicz , Wpływ zasiłku na proces poszukiwania pracy	294
Marta Dziechciarz-Duda, Klaudia Przybysz , Wykształcenie a potrzeby rynku pracy. Klasyfikacja absolwentów wyższych uczelni.....	303
Tomasz Klimanek , Problem pomiaru procesu dezagrarnizacji wsi polskiej w świetle wielowymiarowych metod statystycznych.....	313
Małgorzata Sej-Kolasa, Mirosława Sztemberg-Lewandowska , Wybrane metody analizy danych wzdluznych.....	321
Artur Zaborski , Zastosowanie miar odległości dla danych porządkowych do agregacji preferencji indywidualnych	330
Mariola Chrzanowska, Nina Drejerska, Iwona Pomianek , Zastosowanie analizy korespondencji do badania sytuacji mieszkańców strefy podmiejskiej Warszawy na rynku pracy.....	338

Katarzyna Wawrzyniak , Klasyfikacja województw według stopnia realizacji priorytetów Strategii Rozwoju Kraju 2007-2015 z wykorzystaniem wartości centrum wierszowego	346
---	-----

Summaries

Eugeniusz Gatnar , Statystyka bilansu płatniczego a konkurencyjność gospodarki Polski	22
Andrzej Sokółowski, Magdalena Czaja , Cluster separability and the effectiveness of k -means method	29
Barbara Pawelek, Józef Pocięcha, Adam Sagan , Multisectoral analysis of latent transitions in bankruptcy prediction models.....	38
Elżbieta Golata , Differences in the process of aging and demographic structures in Poznań and the agglomeration compared to selected Polish cities in the years 2002-2011	48
Aleksandra Łuczak, Feliks Wysocki , Determination of weights for features in problems of linear ordering of objects	59
Marek Walesiak , Reinforcing measurement scale for ordinal data in multivariate statistical analysis	68
Paweł Lula , Automatic identification of keywords and keyphrases in documents written in Polish.....	76
Mariusz Kubus , The proposition of modification of the relaxed LASSO method.....	84
Andrzej Bąk, Tomasz Bartłomowicz , Microeconomic multinomial logit models and their implementation in the <code>DiscreteChoice</code> R package .	94
Justyna Brzezińska , The analysis of unemployment data in Poland in 2004-2012 with application of log-linear models	103
Andrzej Bąk, Marcin Pelka, Aneta Rybicka , Application of the MMLM package of R software for vodka consumers preference analysis.....	112
Barbara Batóg, Jacek Batóg , Analysis of the stability of classification of Polish voivodeships in 2002-2010 according to the sectoral labour productivity	120
Małgorzata Markowska, Danuta Strahl , Classification of the European regional space in terms of smart growth pillars using the reference limit system.....	130
Kamila Migdał Najman, Krzysztof Najman , Formal quality assessment of group structure mapping on the Kohonen's map	138
Kamila Migdał Najman, Krzysztof Najman , Graphical quality assessment of group structure mapping on the Kohonen's map	147
Beata Basiura, Anna Czapkiewicz , Validation of time series clustering	156
Michał Trzęsiok , Selected methods for outlier detection.....	166

Grażyna Dehnel, Tomasz Klimanek , Taxonomic aspects of indirect estimation accounting for spatial correlation in enterprise statistics	176
Michał Bernard Pietrzak, Justyna Wilk , Economic distance in modeling spatial phenomena with the application of gravity model.....	185
Maciej Beręsewicz , An attempt to use different distance measures in the Generalized Petersen estimator	194
Marcin Szymkowiak, Tomasz Józefowski , Construction and practical using of SPREE estimators for two-dimensional contingency tables.....	201
Marcin Pelka , The ensemble conceptual clustering for symbolic data.....	209
Małgorzata Machowska-Szewczyk , Evaluation of clusters obtained by fuzzy classification methods for symbolic objects.....	219
Justyna Wilk , Problem of determining the number of clusters in taxonomic analysis of symbolic data	228
Andrzej Dudek , Clustering techniques for Google maps markers.....	236
Ewa Roszkowska , The evaluation of negotiation offers in ill structure negotiation problems with the application of fuzzy SAW procedure	247
Marcin Szymkowiak, Marek Witkowski , The use of correspondence analysis in analysing the financial situation of cooperative banks.....	256
Bartłomiej Jefmański , The construction of fuzzy customer satisfaction indexes using R program.....	265
Karolina Bartos , Discovering patterns of consumer behaviour by market basket analysis of the transactional data.....	274
Joanna Trzęsiok , Cluster analysis of countries with respect to fertility rate and other demographic factors	284
Beata Bal-Domańska , An attempt to identify major regional clusters and their convergence	293
Beata Bieszk-Stolorz, Iwona Markowicz , The influence of benefit on the job finding process	302
Marta Dziechciarz-Duda, Klaudia Przybysz , Education and labor market needs. Classification of university graduates	312
Tomasz Klimanek , The problem of measuring deagrarianisation process in rural areas in Poland using multivariate statistical methods.....	320
Małgorzata Sej-Kolasa, Mirosława Sztemberg-Lewandowska , Selected methods for an analysis of longitudinal data.....	329
Artur Zaborski , The application of distance measures for ordinal data for aggregation individual preferences	337
Mariola Chrzanowska, Nina Drejerska, Iwona Pomianek , Application of correspondence analysis to examine the situation of the inhabitants of Warsaw suburban area in the labour market	345
Katarzyna Wawrzyniak , Classification of voivodeships according to the level of the realization of priorities of <i>the National Development Strategy 2007-2015</i> with using the values of centroid of the rows	355

Bartłomiej Jefmański

Uniwersytet Ekonomiczny we Wrocławiu

BUDOWA ROZMYTYCH INDEKSÓW SATYSFAKCJI KLIENTÓW Z ZASTOSOWANIEM PROGRAMU R

Streszczenie: W artykule scharakteryzowano wybrane funkcje i pakiety programu R, które mogą znaleźć zastosowanie w budowie indeksów satysfakcji klientów z użyciem liczb rozmytych. Rozważania dotyczą pewnej klasy indeksów satysfakcji, budowanych na podstawie subiektywnych modeli oceny jakości usług. Przykład empiryczny umożliwił wizualizację statystyk opisowych dla liczb rozmytych istotnych z punktu widzenia konstrukcji indeksów satysfakcji.

Słowa kluczowe: indeksy satysfakcji, statystyki liczb rozmytych, program R.

1. Wstęp

Proces budowy indeksów satysfakcji klienta jest złożony i na różnych etapach wykorzystuje różne metody z dorobku wielowymiarowej analizy statystycznej. Niezależnie jednak od wybranego schematu postępowania oraz stosowanych metod statystycznych w ramach poszczególnych etapów konstrukcji indeksu nowym trendem w ich budowie jest łączenie tych metod z dorobkiem teorii zbiorów rozmytych, a w szczególności z pewną klasą tych zbiorów, określaną mianem liczb rozmytych. Nie jest to jedyne zastosowanie zbiorów rozmytych w obszarze badań satysfakcji z usług, ale obecnie dominuje w literaturze przedmiotu. Główną przesłanką podejmowanych badań w tym obszarze jest nieprecyzyjność informacji, na podstawie których szacowane są wartości indeksów. Źródło tej nieprecyzyjności tkwi w sposobie pomiaru opinii respondentów, a ściślej – w niejednoznaczności kategorii stosowanych do opisu punktów skal szacunkowych stosowanych w badaniach satysfakcji. Transformacja tych kategorii do postaci liczb rozmytych jest jednym z możliwych rozwiązań.

Konstrukcja indeksów satysfakcji z zastosowaniem liczb rozmytych, a następnie analiza uzyskanych w ten sposób wyników wymaga odpowiedniego oprogramowania statystycznego. Jednym z nielicznych programów oferujących działania

na tych liczbach oraz obliczanie statystyk opisowych niezbędnych w budowie indeksów satysfakcji jest program R. Dlatego zasadniczym celem artykułu jest przegląd oraz charakterystyka funkcji i pakietów programu R, które mogą znaleźć zastosowanie w budowie indeksów satysfakcji klientów. Artykuł adresowany jest do badaczy zajmujących się tematyką badania satysfakcji klientów z zastosowaniem indeksów satysfakcji i zbiorów rozmytych. Zawarte w artykule składnie poleceń mają umożliwić budowę tego typu narzędzi w programie R bez konieczności biegłej znajomości tego oprogramowania.

2. Charakterystyka funkcji programu R przydatnych w budowie rozmytych indeksów satysfakcji klienta

Można wymienić wiele podejść w konstrukcji indeksów satysfakcji z usług. Dużą popularnością cieszą się indeksy satysfakcji oparte na modelach subiektywnej oceny jakości usług (np. indeks SERVQUAL, indeks WebQual lub CSI indeks). Obecnie metodologia tych indeksów rozwijana jest w połączeniu z analizą liczb rozmytych. Przykłady takich zastosowań można znaleźć m.in. w pracach: [Fasanghari, Roudsari 2008; Liu i in. 2008; Carrasco i in. 2012; Zani i in. 2010, 2012; Erdoğan i in. 2013].

W programie R dostępne są trzy pakiety, które mogą znaleźć zastosowanie w budowie indeksów satysfakcji opartych na liczbach rozmytych: `fuzzyOP` [Aklan i in. 2008], `FuzzyNumbers` [Gagolewski 2013], `SAFD` [Trutschnig i in. 2013]. W pakietach stosowane są funkcje umożliwiające generowanie statystyk opisowych dla liczb rozmytych oraz wykonywanie operacji arytmetycznych na tych liczbach, które zostały opisane m.in. w opracowaniu Viertla [2011].

Pierwszy z pakietów posiada zaimplementowane funkcje umożliwiające przeprowadzenie większości operacji arytmetycznych na liczbach rozmytych oraz pozwalających na wizualizację otrzymanych wyników. Umożliwia również obliczenie podstawowych statystyk opisowych.

Pakiet `FuzzyNumbers`, podobnie jak `fuzzyOP`, został opracowany do analizy danych rozmytych w postaci liczb rozmytych. Oferuje możliwość wykonania operacji arytmetycznych na liczbach rozmytych oraz wizualizację wyników w postaci liczb rozmytych. Pozwala obliczyć odległość Euklidesa (lub jej kwadrat) między dwoma liczbami rozmytymi. Posiada również funkcje do wyostrzania liczb rozmytych.

Pakiet `SAFD` oferuje funkcje do statystycznej analizy danych rozmytych w postaci liczb rozmytych. Oprócz działań arytmetycznych na tych liczbach pozwala na obliczenie takich statystyk, jak mediana, wariancja czy współczynnik korelacji. Ma wbudowaną funkcję losowego generowania liczb rozmytych oraz umożliwia szacowanie funkcji regresji dla liczb rozmytych o trapezoidalnej postaci funkcji przynależności. Posiada również funkcję pozwalającą na wyostrzenie wyników analiz otrzymanych w postaci liczb rozmytych.

Z punktu widzenia konstrukcji indeksów satysfakcji z zastosowaniem liczb rozmytych można wyróżnić trzy zasadnicze etapy:

1. transformację wyników pomiaru opinii respondentów do postaci liczb rozmytych i wizualizację otrzymanych wyników,

2. obliczenie wyników działań arytmetycznych oraz statystyk opisowych związanych z ustaleniem systemu wag zmiennych oraz obliczeniem wartości funkcji agregującej,

3. wyostrzenie wartości indeksu w postaci liczby rozmytej.

Funkcje wyszczególnionych powyżej pakietów, które mogą znaleźć zastosowanie na poszczególnych etapach konstrukcji indeksów, scharakteryzowano w tab. 1.

Tabela 1. Funkcje i ich składnie dla poszczególnych etapów konstrukcji indeksów satysfakcji z zastosowaniem liczb rozmytych

Etap konstrukcji indeksu		Funkcje		
		pakiet fuzzyOP	pakiet SAFD	pakiet FuzzyNumbers
Transformacja danych do postaci liczb rozmytych oraz wizualizacja wyników		fuzzynumber(anz, A, wektor, pl)	–	TriangularFuzzyNumber(a1, amid, a4) TrapezoidalFuzzyNumber(a1, a2, a3, a4)
Ważenie oraz agregacja zmiennych	suma	fuzzysum(anz, A, wektor, pl)	Msum(XX)	operator (+)
	średnia	fuzzymeans(anz, A, wektor, pl)	Mmean(XX, pic)	–
	iloczyn	fuzzyproduct(anz, A, wektor, pl)	–	operator (*)
	mnożenie przez wartość skalarną	fuzzy-scalar(anz, A, wektor, v, pl)	sc_mult(X, b)	–
Wyostrzenie wartości indeksu		–	Defuzzify(XX)	value(object, ...) weightedExpectedValue(object, w, ...)

Źródło: opracowanie własne na podstawie [Aklan i in. 2008; Trutschnig i in. 2013; Gagolewski 2013].

Składnia funkcji pakietu fuzzyOP wyszczególnionych w tabela 1 jest bardzo podobna. Najważniejsze wspólne argumenty tych funkcji to: *anz* – liczba α -przekrojów; *A* – zbiór danych zawierający co najmniej dwie kolumny (pierwsza kolumna prezentuje punkty wsparcia liczby rozmytej, a druga odpowiadające im wartości funkcji charakterystycznej); *wektor* – wektor wskazujący, które liczby rozmyte ze zbioru danych mają zostać uwzględnione w obliczeniach; *v* – wartość skalarna; *pl* – opcja wykresu: 0 – brak wykresu, 1 – wykres funkcji charakterystycznej liczby rozmytej stanowiącej wynik działania funkcji (np. średniej liczby rozmytej w przypadku funkcji *fuzzymeans*, minimalnej liczby rozmytej dla funk-

cji `fuzzymin` itp.), 2 – wykres zbudowany z dwóch części, zawierający funkcje charakterystyczne liczb rozmytych wskazanych przez argument `vektor`, oraz funkcję charakterystyczną wyniku działania wybranej funkcji.

W przypadku funkcji pakietu `SAFD`, mogących znaleźć zastosowanie w budowie indeksów satysfakcji, wyjaśnienia wymagają trzy argumenty:

`XX` – lista liczb rozmytych; `X` – tablica danych zawierająca liczby rozmyte; `pic` – opcja wykresu: 0 – brak wykresu, 1 – wykresy średniej arytmetycznej w postaci liczby rozmytej; `b` – wartość skalarna.

Pakiet `FuzzyNumbers` oferuje kilka sposobów tworzenia i wizualizacji liczb rozmytych. W tabeli 1 wyszczególniono funkcje właściwe dla trójkątnych i trapezoidalnych liczb rozmytych, będących przypadkiem liczb rozmytych, najczęściej stosowanych w budowie indeksów satysfakcji. Wyszczególnione funkcje zawierają następujące argumenty: `a1` – wartość określająca lewą granicę wsparcia liczby rozmytej; `a2` – wartość określająca lewą granicę jądra liczby rozmytej (dla przypadku trapezoidalnej liczby rozmytej); `a3` – wartość określająca prawą granicę jądra liczby rozmytej (dla przypadku trapezoidalnej liczby rozmytej); `a4` – wartość określająca prawą granicę wsparcia liczby rozmytej; `amid` – wartość określająca jądro liczby rozmytej (dla przypadku trójkątnej liczby rozmytej); `object` – liczba rozmyta; `w` – wartość z przedziału $[0; 1]$.

3. Przykład zastosowania wybranych funkcji dla liczb rozmytych w konstrukcji indeksu WebQual

Indeks WebQual to narzędzie do oceny jakości stron internetowych. Jego rozmyta modyfikacja polega na zastosowaniu w ocenie ważności oraz percepcji kryteriów strony internetowej liczb rozmytych. W niniejszym przykładzie zaprezentowane zostaną wybrane funkcje pakietu `fuzzyOP` oraz `FuzzyNumbers`, umożliwiające oszacowanie wartości indeksu według schematu zaproponowanego w opracowaniu Jefmańskiego [2013]. Dane wejściowe miały formę trójkątnych liczb rozmytych przyporządkowanych poszczególnym punktom dwóch szacunkowych skal pomiarowych. Respondenci oceniali ważność oraz percepcję każdego z kryteriów za pomocą skal pięciostopniowych (z opisanymi wszystkimi punktami skali w postaci wartości lingwistycznych). W przypadku oceny ważności do wyboru była jedna z pięciu kategorii: zdecydowanie nieważne, nieważne, średnio ważne, ważne, zdecydowanie ważne. W ocenie percepcji zastosowano następujące kategorie: bardzo niska, niska, średnia, wysoka, bardzo wysoka. Dla każdej z wyszczególnionych kategorii respondenci określili wartość z przedziału $\langle 0\%;100\% \rangle$, która ich zdaniem najlepiej oddaje natężenie danej kategorii. Kategorie zostały następnie transformowane do postaci trójkątnych liczb rozmytych, których zakresy dziedzin unormowano w przedziale $\langle 0;100 \rangle$. Trzy parametry a , b i c dla każdej z liczb rozmytych zostały oszacowane odpowiednio na podstawie wartości: minimalnej,

dominującej oraz maksymalnej, obliczonych z próby badawczej. Parametry liczb rozmytych miały następujące wartości: „zdecydowanie nieważne” (0; 10; 30), „nieważne” (0; 30; 50), „średnio ważne” (5; 50; 75), „ważne” (10; 80; 90), „zdecydowanie ważne” (60; 90; 100), „bardzo niska” (0; 10; 40), „niska” (0; 30; 60), „średnia” (5; 50; 80), „wysoka” (15; 80; 90), „bardzo wysoka” (50; 90; 100).

W niniejszym przykładzie zaprezentowany zostanie sposób szacowania indeksu dla pierwszego z 18 kryteriów jakości strony internetowej. Zgodnie z założeniem metody, pozostałe indeksy szacowane są w identyczny sposób, a index WebQual dla strony internetowej stanowi uśrednienie wartości indeksów dla wszystkich kryteriów.

W analizie wykorzystano dwa zbiory danych zawierające odpowiednio oceny ważności oraz percepcji 102 respondentów. Każdy zbiór miał postać macierzy danych, gdzie dwie kolejne kolumny macierzy przedstawiają odpowiednio punkty wsparcia liczby rozmytej oraz wartość funkcji charakterystycznej dla tych punktów.

Konstrukcję indeksu należy rozpocząć od obliczenia dla każdego kryterium oceny ważności, którą stanowi uśredniony wynik z wszystkich ocen respondentów. W tym celu można zastosować funkcję `fuzzymean`, która pozwoli obliczyć średnią ocenę ważności w postaci liczby rozmytej. W wyniku zastosowania poniższego skryptu otrzymuje się średnią ocenę ważności pierwszego kryterium (`k1_w`):

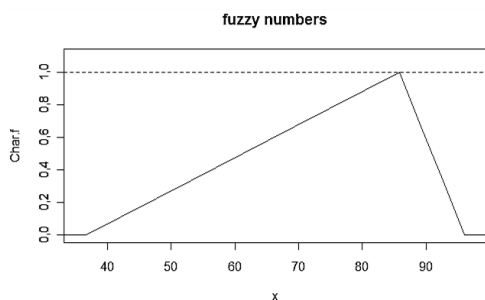
```
# Ocena ważności pierwszego kryterium jakości
library(fuzzyOP)
x <- read.csv2("dane_waznosc.csv", header=FALSE)
options(OutDec=",")
k1_w<-fuzzymean(3, x, c(1:204), 1)
```

Średnia ocena ważności ma postać liczby rozmytej o następujących parametrach: $a_1 = 36,65$; $a_2 = 85,85$; $a_3 = 96,10$. W podobny sposób można obliczyć średnią ocenę postrzegania jakości pierwszego kryterium (`k1_p`):

```
# Ocena percepcji pierwszego kryterium jakości
library(fuzzyOP)
y <- read.csv2("dane_percepcja.csv", header=FALSE)
options(OutDec=",")
k1_p<-fuzzymean(3, x, c(1:204), 1)
```

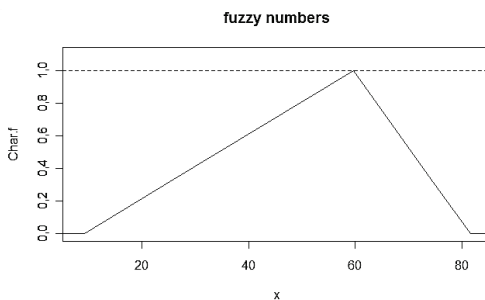
Średnia ocena percepcji również wyrażona jest w postaci liczby rozmytej o następujących parametrach: $a_1 = 9,21$; $a_2 = 59,63$; $a_3 = 81,59$. Powyższe skrypty generują ocenę ważności oraz percepcji w postaci trójkątnych liczb rozmytych przedstawionych odpowiednio na rys. 1 i 2.

W kolejnym etapie budowy indeksu obliczana jest ważona ocena każdego kryterium. W tym celu można zastosować funkcję `weightedExpectedValue` pakietu `FuzzyNumbers` oraz `fuzzyscalar` pakietu `fuzzyOP`. Wartości



Rys. 1. Liczba rozmyta reprezentująca średnią ocenę ważności pierwszego kryterium

Źródło: opracowanie własne z zastosowaniem pakietu `fuzzyOP`.



Rys. 2. Liczba rozmyta reprezentująca średnią ocenę postrzegania pierwszego kryterium

Źródło: opracowanie własne z zastosowaniem pakietu `fuzzyOP`.

wyostrzone liczb rozmytych otrzymuje się zgodnie z metodą ważonej wartości oczekiwanej zaproponowaną przez Dubois i Prade'a [1987]. W wyniku zastosowania poniższego skryptu otrzymuje się ocenę ważoną pierwszego kryterium:

```
# Obliczenie wyostrzonej oceny pierwszego kryterium
jakości
library(FuzzyNumbers)
k1_p<-TriangularFuzzyNumber(9.21, 59.63, 81.59)
k1_p_def<-weightedExpectedValue(k1_p, w=0.5)

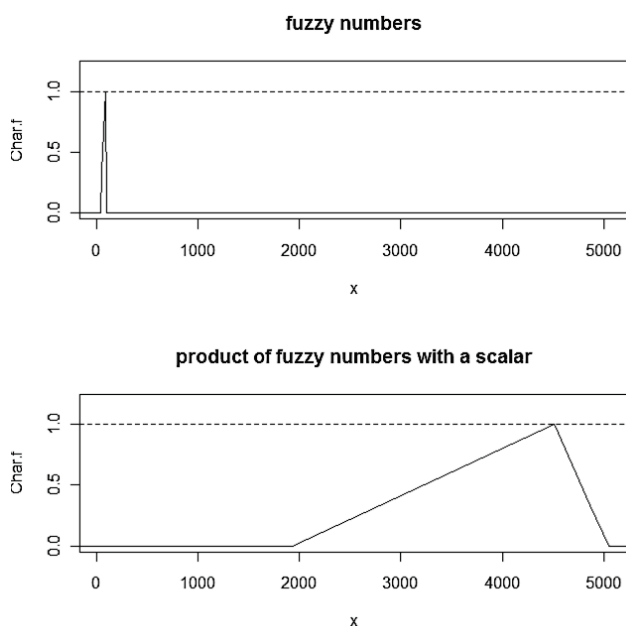
# Ocena ważona pierwszego kryterium jakości
library(fuzzyOP)
a<-c(36.65, 85.85, 96.10)
b<-c(0, 1, 0)
k1_w<-cbind(a, b)
k1_p_w<-fuzzyscalar(3, k1_w, 1, 52.52, 2)
```

Wyostrzona ocena ważności pierwszego kryterium wyniosła 52,52. Ocena ważona tego kryterium ma postać trójkątnej liczby rozmytej o następujących parametrach: $a_1 = 1924,86$; $a_2 = 4508,84$; $a_3 = 5047,17$. Wizualizację wyników otrzymanych z zastosowaniem powyższego skryptu przedstawiono na rys. 3.

Wartość indeksu WebQual dla j -tego kryterium ustalana jest jako iloraz ważonej oceny j -tego kryterium oraz maksymalnej (możliwej do uzyskania) oceny j -tego kryterium. W prezentowanym przykładzie maksymalna ocena ważona dla pierwszego kryterium ma postać liczby rozmytej o następujących parametrach: $a_1 = 2931,71$; $a_2 = 6868,29$; $a_3 = 7687,8$. Ocena ważona pierwszego kryterium wyniosła: $a_1 = 1837,53$; $a_2 = 4304,9$; $a_3 = 4818,55$. Obliczenie indeksu WebQual dla pierwszego kryterium, zgodnie z metodą zaprezentowaną w opracowaniu Jefmańskiego [2013], uzyskuje się poprzez wyostrzenie oceny ważonej oraz maksymalnej na co pozwala poniższy skrypt:

```
# Wyostrzenie maksymalnej oceny ważonej dla pierwszego
kryterium jakości
library(FuzzyNumbers)
k1_p_w_max<-TriangularFuzzyNumber(2931.71, 6868.29, 7687.80)
k1_p_w_def<-weightedExpectedValue(k1_p_w_max, w=0.5)

# Wyostrzenie ważonej oceny pierwszego kryterium
library(FuzzyNumbers)
k1_p_w<- TriangularFuzzyNumber(1924.86, 4508.84, 5047.17)
k1_p_w_def<-weightedExpectedValue(k1_p_w, w=0.5)
```



Rys. 3. Ocena ważona dla pierwszego kryterium w postaci liczby rozmytej

Źródło: opracowanie własne z zastosowaniem pakietu *fuzzyOP*.

Otrzymano ocenę maksymalną oraz ważoną odpowiednio na poziomie: 6089,02 i 3997,43. Wartość indeksu (wyrażona procentowo) dla pierwszego kryterium wyniosła 65,65%. W ten sam sposób można obliczyć indeksy dla pozostałych kryteriów. Wartość indeksu WebQual dla ocenianej strony internetowej stanowi średnią arytmetyczną z wartości indeksów szacowanych na poziomie poszczególnych kryteriów.

4. Podsumowanie

W przypadku budowy rozmytych indeksów satysfakcji klientów, konstruowanych na podstawie modeli oceny jakości usług, zastosowanie znajdują przede wszystkim pakiety `fuzzyOP` i `FuzzyNumbers`. Umożliwiają one analizę liczb rozmytych (w tym liczb trójkątnych i trapezoidalnych), które w przeważającej części opracowań poświęconych budowie rozmytych indeksów satysfakcji klientów stanowią podstawę transformacji wyników pomiaru ze skali porządkowej.

Nie można wskazać „kompletnego” pakietu z punktu widzenia budowy rozmytych indeksów satysfakcji. Słabością pakietu `fuzzyOP` jest brak możliwości wyostrzania liczb rozmytych – co jest istotne z punktu widzenia możliwości porównywania respondentów lub organizacji pod względem wartości rozmytego indeksu. Lukę tę wypełnia jednak pakiet `FuzzyNumbers`, który oferuje wyostrzanie liczb rozmytych z zastosowaniem dwóch metod. Niemniej jednak wybór metod wyostrzania liczb rozmytych w omówionych pakietach jest stosunkowo mały.

Funkcje dostępne w ramach pakietów `fuzzyOP` i `FuzzyNumbers` można również wykorzystać w wybranych etapach konstrukcji pozostałych typów indeksów satysfakcji klienta. W przypadku budowy kompozytowych indeksów satysfakcji można połączyć wyniki metod wielowymiarowej analizy statystycznej (zwłaszcza metod eksploracyjnych) z analizą liczb rozmytych, np. na etapie pomiaru opinii respondentów. W przypadku budowy indeksów opartych na metodach porządkowania liniowego zastosowanie znajdują przede wszystkim pakiety `SAFD` i `FuzzyNumbers`, które pozwalają obliczyć odległość (np. euklidesową) między dwoma liczbami rozmytymi.

Literatura

- Akkan S., Altindas E., Macit R., Umar S., Unal H. (2008), *FuzzyOP: Fuzzy Numbers and the Main Mathematical Operations*, <http://CRAN.R-project.org/package=fuzzyOP>.
- Carrasco R.A., Munoz-Leiva F., Sánchez-Fernández J., Liébana-Cabanillas F.J. (2012), *A model for the integration of e-financial services questionnaires with SERVQUAL scales under fuzzy linguistic modeling*, „Expert Systems with Applications”, vol. 39, s. 1535-1547.
- Dubois D., Prade H. (1987), *The mean value of a fuzzy number*, „Fuzzy Sets and Systems”, vol. 24, s. 279-300.
- Erdoğan M., Bilişik Ö.N., Kaya İ., Baraç H. (2013), *A customer satisfaction model based on fuzzy TOPSIS and SERVQUAL methods*, „Lecture Notes in Management Science”, vol. 5, s. 74-83.
- Fasanghari M., Roudsari F.H. (2008), *The Fuzzy Evaluation of e-Commerce Customer Satisfaction*, „Word Applied Sciences Journal”, vol. 4, no. 2, s. 164-168.
- Gagolewski M. (2013), *Tools to deal with fuzzy numbers*, <http://www.rexamine.com/resources/fuzzynumbers> (30.09.2013).

- Jefmański B. (2013), *Rozmyta metoda WebQual i jej zastosowanie w ocenie internetowych serwisów informacyjnych*, „Handel Wewnętrzny” (w druku).
- Liu X., Zeng X., Xu Y., Koehl L. (2008), *A fuzzy model of customer satisfaction index in e-commerce*, „Mathematics and Computers in Simulation”, vol. 77, iss. 5-6, s. 512-521.
- Trutschnig W., Lubiano A., Lastra J. (2013), *SAFD – An R Package for Statistical Analysis of Fuzzy Data*, „Studies in Fuzziness and Soft Computing”, vol. 285, s. 107-118.
- Viertl R. (2011), *Statistical Methods for Fuzzy Data*, Wiley & Sons, Chichester.
- Zani S., Milioli A., Morlini I. (2010), *Fuzzy Composite Indicators: An Application for Measuring Customer Satisfaction*, [w:] N. Torelli, F. Pesarin, A. Bar-Hen (red.), *Advances in Theoretical and Applied Statistics*, Springer, Berlin – Heidelberg.
- Zani S., Milioli A., Morlini I. (2012), *Fuzzy Methods and Satisfaction Indices*, [w:] R.S. Kenett, S. Salini (red.), *Modern Analysis of Customer Surveys with applications using R*, Wiley & Sons, New York.

THE CONSTRUCTION OF FUZZY CUSTOMER SATISFACTION INDEXES USING R PROGRAM

Summary: The paper describes selected functions and packages of R program which can be used in the construction of customer satisfaction indexes with an application of fuzzy numbers. The considerations were related to a class of satisfaction indexes built on the basis of subjective models of service quality evaluation. The empirical example allowed for a visualization of descriptive statistics for fuzzy numbers relevant to the construction of satisfaction indexes.

Keywords: satisfaction indexes, statistics for fuzzy numbers, R program.